CH5350: Applied Time-Series Analysis

Auto-Correlation Functions

Arun K. Tangirala

Department of Chemical Engineering, IIT Madras

Opening remarks

A prime goal of TSA is **prediction** (forecasting). Therefore, it is important to test for predictability of a series prior to any model development exercise.

Predictability is a generic term and can largely depend on what model is being considered for prediction. Most of the established theory revolves around **linear models** for predictions, which serve a large number of applications.

Recall that covariance (or correlation) is a measure of linear dependence between two random variables. **We shall now apply this concept to two observations of a series so as to test for linear dependence within that series.**

Auto-covariance function (ACVF)

The **auto-covariance function** (ACVF) is defined as the covariance between two observations of a series, $v[k_1]$ and $v[k_2]$

$$\boxed{\sigma_{vv}[k_1, k_2] = E((v[k_1] - \mu_{k_1})(v[k_2] - \mu_{k_2}))} \tag{1}$$

where μ_{k_i} is the mean of the process at k_i instant.

Note: We have switched our notation from $x[k]$ to $v[k]$ and used $v[k_i]$ to indicate both the observation as well as the associated RV.

ACVF of stationary processes

For stationary processes, recall that the mean remains invariant and the distribution is only a function of the time difference or lag, $l = k_1 - k_2$. Consequently,

ACVF of a stationary process

The auto-covariance function of a stationary process is only a function of the **lag** l between two observations,

$$\boxed{\sigma_{vv}[l] = E((v[k] - \mu_v)(v[k-l] - \mu_v))} \tag{2}$$

where $\mu_v = E(v[k])$ is the mean of the stationary process

Properties of the ACVF

- $\sigma_{vv}[l]$ **measures (only) the linear dependence** between $v[k]$ and $v[k-l]$.
- By virtue of stationarity, **ACVF is a symmetric measure**, *i.e.*,

$$\boxed{\sigma_{vv}[l] = \sigma_{vv}[-l]} \tag{3}$$

- It lacks directionality, *i.e.*, $\sigma_{vv}[l]$ does not provide the direction of dependence
- It is affected by confounding, *i.e.*, $\sigma_{vv}[l]$ includes the effects of other observations that can commonly influence $v[k]$ and $v[k-l]$
- The value of ACVF depends on the units in which the series is expressed.

Auto-correlation function (ACF)

In order to address the unboundedness and sensitivity to choice of units, the **auto-correlation function** (ACF) is introduced.

$$\boxed{\rho_{vv}[l] = \frac{\sigma_{vv}[l]}{\sigma_{vv}[0]}} \tag{4}$$

Remark: The ACF possesses all characteristics of correlation.

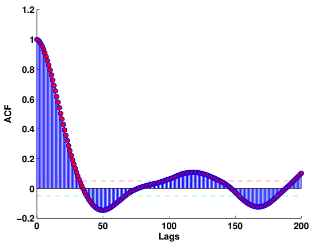
Properties of ACF

- It reaches a maximum value of 1 at lag 0 - the dependency of a sample on itself is normalized to unity.
- In fact, it is bounded like the correlation

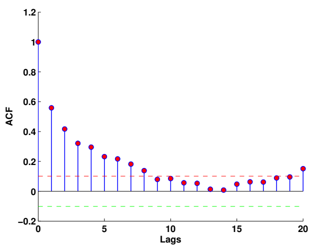
$$-1 \leq \rho_{vv}[l] \leq 1 \tag{5}$$

The equality occurs if and only when $v[k] = \alpha v]k - l], \quad \alpha \in \mathcal{R}$, *i.e.*, for a purely linear deterministic process

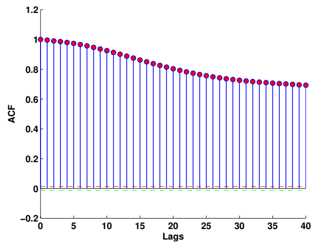
Sample ACFs of some real series



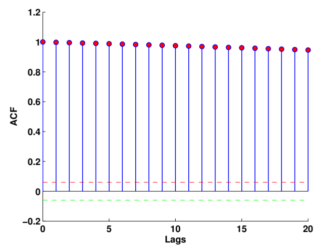
(a) Temperature



(b) Wind speed



(c) ECG



(d) Swiss SMI

$\boxed{\text{The decay rates for ACF depends on the series under analysis.}}$

By examining the ACF we obtain useful insights into the nature of correlation and the type of model that can be possibly built.

Interpreting ACF in predictions

Consider the linear forecast of a series at $k_2 = k + l$ given information only at $k_1 = k$

$$\hat{v}[k + l | v[k]] = \alpha v[k] \quad (6)$$

Then, the optimal value of α in the sense of

$$\min_{\alpha} E\left((v[k + l] - \hat{v}[k + l | v[k]])^2\right) = \min_{\alpha} E\left((v[k + l] - \alpha v[k])^2\right)$$

is

$$\boxed{\alpha^\star = \rho_{vv}[l]}$$

Thus, the ACF at any lag l is the optimal coefficient of the linear model in (6)

Discovering process signatures from ACF

Methods for development / estimating time-series models implicitly or explicitly involve inverse mapping of the estimated ACF to the model parameters.

Therefore, it is important to ensure that this inverse mapping produces mathematically meaningful and correct models. For instance, **can we start from any symmetric function and construct a model of stationary random process?** NO.

Non-negative definiteness

An important property that a symmetric function to serve as the ACF of a stationary random process is that of **non-negative definiteness**.

Definition

A sequence $\zeta[.]$ is said to be non-negative definite if and only if it satisfies

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i \zeta[|i-j|] a_j \geq 0 \qquad \forall \; a_i, a_j \in \mathcal{R} \;, \; \forall n > 0 \tag{7}$$

Alternatively, the sequence is non-negative if and only if the matrices $\Sigma_n = (\zeta[i-j])_{i,j=1}^{n}$ are non-negative definite $\forall n$.

Non-negative definiteness of ACF

Theorem

The ACVF of a stationary process is non-negative definite.

Proof: Consider a process $y[k] = \sum_{i=1}^{n} a_i v[k-i+1] = \mathbf{a}^T\mathbf{v}$, where $v[.]$ is an observation of a random stationary process and $a_i \in \mathcal{R}$. Then

$$\begin{aligned}
\text{var}(y[k]) &= E\left((y[k]-\mu_y)(y[k]-\mu_y)^T\right) \\
&= E\left(\mathbf{a}^T(\mathbf{v}-\mu_{\mathbf{v}})(\mathbf{v}-\mu_{\mathbf{v}})^T\mathbf{a}\right) \\
&= \mathbf{a}^T\Sigma_n\mathbf{a} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\sigma[i-j]a_j \geq 0
\end{aligned}$$

$$\Sigma_n = \begin{bmatrix} \sigma_{vv}[0] & \cdots & \sigma_{vv}[n-1] \\ \vdots & \cdots & \vdots \\ \sigma_{vv}[1-n] & \cdots & \sigma_{vv}[0] \end{bmatrix}$$

White-noise process: Purely (linear) random process

One of the most important uses of ACF is in the definition of an **ideal random process**, which is the backbone of (linear) random process theory.

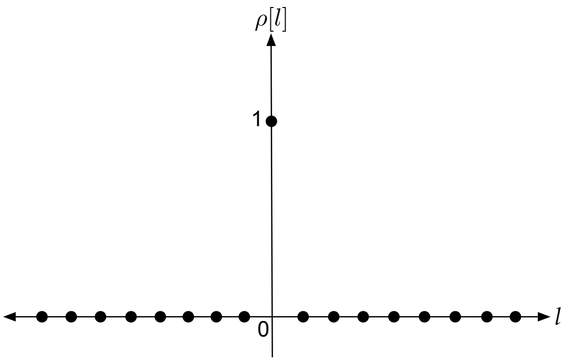
White-noise process

The white-noise process $e[k]$ is a **stationary uncorrelated** random process,

$$\rho_{ee}[l] = \begin{cases} 1 & l = 0 \\ 0 & l \neq 0 \end{cases} \tag{8}$$

White-noise process . . . contd.

- It is an unpredictable (in the linear sense) stationary process.
- The ACF of a white-noise process has an impulse-like shape. For any process with predictability, the ACF deviates from this shape.



Role of WN process in TSA

The white-noise (WN) process is useful in two important ways:

1. Serves as the **benchmark for (lack of) predictability** at two different stages of time-series modelling:
   - Pre-modelling stage: Testing the **given series**
   - Post-modelling stage: Whiteness test of **residuals**
2. As a **fictitious input** (driving force) to a random process for modelling purposes

Observe that the definition of WN does not impose any conditions on the distribution - the only requirements are **stationarity** and **uncorrelated** properties

WN processes

In principle, therefore we can conceive Gaussian WN, Uniform WN and so on. The most commonly assumed one is the Gaussian WN (GWN).

Remark: A variety of random number generators can generate GWN and UWN processes. These are pseudo-random number generators in the sense that they lose their randomness when the initial condition (seed) is known.

Independence and the I.I.D. (ideal random) process

One can extend the definition of uncorrelated process to an independent process, which demands that all higher-order moments of the joint pdf to be zero.

I.I.D. process

An identical, independent process is that process which is absolutely unpredictable (using any non-linear model).

- **Note:** A Gaussian white-noise process is an i.i.d process as well.
- In practice, it is very difficult to test for independence whereas it is quite easy to test for the absence of correlation.

ACF and time-series modelling

To quickly recap, ACF:

- Provides means of testing for predictability in a series.
- Facilitates definition of the ideal random process (white-noise process).

Taking a step further, we now move to the stage of **using ACF for determining what type (order and structure) of linear model is suitable for a given series**. To be able to do so, it is important to study the behaviour of ACF for different processes.

Using the map between ACF and the associated random process, we would like to make an "intelligent" guess of the form of the model.

General linear random process

The general model for a linear random stationary process is given by the **convolution form**

$$v[k] = \sum_{n=-\infty}^{\infty} h[n]e[k-n], \qquad e[k] \sim \text{GWN}(0, \sigma_e^2), \sum_n |h[n]| < \infty \tag{9}$$

- Typically one sets $n \geq 0$ (for **causality**) and $h[0] = 1$ (for **uniqueness**).
- The absolute convergence criterion implies that $h[n] \longrightarrow 0$ as $n \longrightarrow \infty$.

A formal treatment of the above model appears later.

Types of linear random processes

Depending on what one assumes further on how $h[n]$ decays, the linear random process in (9) specializes to two different sub-classes of processes:

1. **Moving Average** (MA) process: $h[n]$ goes to zero **abruptly** after finite n.
2. **Auto-regressive** (AR) process: $h[n]$ decays to zero **asymptotically** with n.

and mixed models, i.e, auto-regressive, moving average (ARMA) models.

We shall now study the ACF signatures of the above two processes.

ACF of a Moving Average (MA) process

MA(1) process

The **MA process of first-order**, i.e., MA(1) arises when $h[1] = c_1$ and $h[n] = 0, n \neq 0, 1$.

$$v[k] = e[k] + c_1 e[k-1]$$

where $e[k] \sim \text{GWN}(0, \sigma_e^2)$ and c_1 is a finite constant.

- The current state contains the previous shock wave plus an unpredictable part $e[k]$

Theoretical ACF of an MA(1) process

The theoretical ACF is obtained using the definition in (2)

$$\begin{aligned}\sigma_{vv}[l] &= E((v[k]-\mu_v)(v[k-l]-\mu_v)) \\ &= E((e[k]+c_1e[k-1])(e[k-l]+c_1e[k-l-1])) \\ &= E(e[k]e[k-l]) + c_1E(e[k]e[k-l-1]) \\ &\quad + c_1E(e[k-1]e[k-l]) + c_1^2E(e[k-1]e[k-l-1]) \\ &= \sigma_{ee}[l] + c_1\sigma_{ee}[l+1] + c_1\sigma_{ee}[l-1] + c_1^2\sigma_{ee}[l]\end{aligned}$$

ACVF of an MA(1) process ...contd.

Using the definition of WN process in (8), the ACVF of the MA(1) process is

$$\boxed{\sigma_{vv}[l] = \begin{cases} (1+c_1^2)\sigma_e^2 & l=0 \\ c_1\sigma_e^2 & l=\pm 1 \\ 0 & l=\pm 2, \pm 3, \cdots \end{cases}} \tag{10}$$

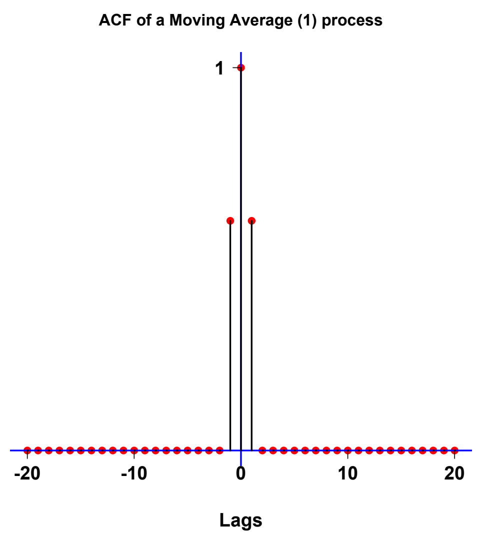
ACF of an MA(1) process ... contd.

Thus, we can write the ACF of an MA(1) process as

$$\boxed{\rho_{vv}[l] = \begin{cases} 1 & l = 0 \\ \dfrac{c_1}{(1+c_1^2)} & l = \pm 1 \\ 0 & |l| \geq 2 \end{cases}} \quad (11)$$

The ACF of an MA(1) process has a sharp cut-off after lag $l = 1$ (the order of the MA(1) process)

ACF of an MA(1) process ... contd.



- Observe that the ACF is independent of the variance of the WN process
- Since $\rho_{vv}[l] = 0, |l| \geq 2$, one cannot predict the process beyond one time-step.
- The ACF is symmetric and bounded above in magnitude by unity for all values of c_1 (verify).

Modelling viewpoint

ACF plays a crucial role in time-series modelling in the sense that models are built from the covariance functions. We shall briefly visit two aspects, that of recovering **unique** and **real-valued** coefficients.

- Suppose that the coefficient on $e[k]$ was c_0 instead of unity. The ACVF is then given by

$$\sigma_{vv}[l] = \begin{cases} (c_0^2 + c_1^2)\sigma_e^2, & l = 0 \\ c_1 c_0 \sigma_e^2, & |l| = 1 \\ 0, & |l| \geq 2 \end{cases} \tag{12}$$

Modelling viewpoint: ...contd.

From a modelling viewpoint, we have two equations and three unknowns (c_0, c_1 and σ_e^2). We have, therefore, an underdetermined problem.

Thus, in order to obtain **unique estimates**, it is required to fix one of these unknowns. A reasonable and meaningful choice is to set $c_0 = 1$ (why?)

We turn to the second aspect next.

What is the guarantee that we obtain a model with real-valued coefficients?

Not any symmetric, bounded function can be used as the ACF for building a time-series model.

Non-negative definiteness: Revisited

Non-negative definiteness of a symmetric, bounded sequence $\sigma[.]$ guarantees the existence of a stationary random process with $\sigma[.]$ as its ACVF.

Example

Problem: Suppose the symmetric function $f[l] = \begin{cases} 1, & l = 0 \\ \alpha, & l = \pm 1 \\ 0, & \text{otherwise} \end{cases}$

is claimed to be the ACF of a stationary process. Are all values of α admissible for time-series modelling?

Modelling viewpoint: Example ... contd.

The question can be answered in many ways. Let us take the first route.

Assuming that $f[l]$ is indeed the ACF of a stationary process, it is clear that MA(1) is best suited for the process. Then, the estimate of c_1 is given by equating the theoretical ACF of MA(1) with the given $f[l]$:

$$\frac{c_1}{1 + c_1^2} = \alpha \Longrightarrow c_1 = \frac{1 \pm \sqrt{1 - 4\alpha^2}}{\alpha}$$

To obtain real-valued roots, it is clear that $1 - 4\alpha^2 \geq 0$, implying

$$|\alpha| \leq \frac{1}{2} \tag{13}$$

Alternative method for testing n.n.d.

Theorem (Bochner, Herglotz)

Any absolutely summable real-valued sequence $\sigma[l],\ l \in \mathcal{Z}$ is non-negative definite if and only if its Fourier transform

$$\gamma(\omega) = \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} \sigma[l] e^{-j\omega l} \tag{14}$$

is non-negative valued at all ω, i.e., $\gamma(\omega) \geq 0, \forall \omega$.

- The quantity $\gamma(\omega)$, as we shall learn later, is called the **spectral density** when $\sigma[.]$ is the ACVF (explains why WN is termed white!).
- It is sufficient to restrict the frequency range to $[0, 2\pi)$ (why?).]

Example ... contd.

Applying B-H theorem to the given sequence $f[l]$, we compute $\gamma(\omega)$ using (14) to obtain

$$\gamma(\omega) = \frac{1}{2\pi}(1 + 2\alpha \cos \omega) \tag{15}$$

which is non-negative $\forall \omega$ if and only if

$$|\alpha| \geq \frac{1}{2}$$

The non-negative definiteness of ACF sequence is therefore necessary to be able to construct models with real-valued coefficients.

ACF of a general MA(M) process

To compute the ACVF of a general MA(M) process, it is convenient to introduce the **auto-covariance generating function**.

For this purpose, we first introduce the **backward shift operator** q^{-1}, such that

$$q^{-1}v[k] = v[k-1] \qquad \text{and} \qquad qv[k] = v[k+1] \tag{16}$$

Note: q^{-1} is an operator and **not** a variable. q is the **forward shift operator.**

ACVGF

Then, one could write the general linear model in (9) as

$$v[k] = \sum_{n=-\infty}^{\infty} h[n]q^{-n}e[k] = H(q^{-1}) \tag{17}$$

$$\text{where} \quad \boxed{H(q^{-1}) = \sum_{n=-\infty}^{\infty} h[n]q^{-n}} \tag{18}$$

is known as the **transfer function operator** and

$$\boxed{H(z^{-1}) = \sum_{n=-\infty}^{\infty} h[n]z^{-n}} \tag{19}$$

the **transfer function.**

Auto-covariance generating function

ACVGF

The auto-covariance generating function is defined as

$$\boxed{g_\sigma(z) = \sum_{l=-\infty}^{\infty} \sigma_{vv}[l] z^{-l}} \tag{20}$$

where z is a variable.

ACVGF ...contd.

The key use of this ACVF generating function stems from the fact that it can be computed directly from the MA representation of the random process.

$$v[k] = H(q^{-1})e[k] \tag{21}$$

$$\implies \boxed{g_\sigma(z) = \sigma_e^2 H(z^{-1})H(z)} \tag{22}$$

where $H(z^{-1})$ is the transfer function introduced in (19).

Example: ACVF of an MA(2) process

Problem: Compute the ACVF of an MA(2) process

$$v[k] = e[k] + c_1 e[k-1] + c_2 e[k-2]$$

Solution: First observe that

$$H(q^{-1}) = 1 + c_1 q^{-1} + c_2 q^{-2}$$

To compute the ACVF, construct the ACVGF by computing the product

$$\begin{aligned} g_\sigma(z) &= \sigma_e^2 H(z^{-1}) H(z) = \sigma_e^2 (1 + c_1 z^{-1} + c_2 z^{-2})(1 + h_1 z + h_2 z^2) \\ &= \sigma_e^2 (c_2 z^{-2} + (c_1 + c_1 c_2) z^{-1} + (1 + c_1^2 + c_2^2) + (c_1 + c_1 c_2) z + c_2 z^2) \end{aligned}$$

ACVGF of an MA(2) process

Comparing with equation (20) and reading off the coefficients of z^{-l}, we obtain

$$\sigma_{vv}[l] = \begin{cases} (1 + c_1^2 + c_2^2)\sigma_e^2, & l = 0 \\ (c_1 + c_1 c_2)\sigma_e^2, & l = 1 \\ c_2 \sigma_e^2, & l = 2 \\ 0, & |l| \geq 3 \end{cases} \tag{23}$$

Thus, as expected, the ACVF of an MA(2) process vanishes at all lags $|l| > 2$ □

Auto-Regressive (AR) processes: ACF

The second class of processes that we consider are the **auto-regressive (AR) processes**

For illustration, consider a first-order, i.e., AR(1) process:

$$v[k] = -d_1 v[k-1] + e[k] \tag{24}$$

where $e[k]$ is the zero-mean GWN process of variance σ_e^2 and d_1 is a finite constant.

- The current state is a linear function of the past state plus the unpredictable $e[k]$
- Assume $|d_1| < 1$ (a condition required for stationarity of $v[k]$)

ACF of an AR(1) process ...contd.

The theoretical ACF can be now obtained using the definition in (2)

Observe that $\mu_e = 0 \implies \mu_v = 0$.

$$\begin{aligned}\sigma_{vv}[l] &= E(v[k]v[k-l]) \\ &= -d_1 E(v[k-1]v[k-l]) + E(e[k]v[k-l]) \\ &= -d_1 \sigma_{vv}[l-1] + \sigma_{ev}[l]\end{aligned}$$

where $\sigma_{ev}[l]$ is the cross-covariance function, *i.e.*, the covariance between $e[k]$ and $v[k-l]$.

ACF of an AR(1) process ...contd.

- By symmetry property of $\sigma_{vv}[l]$, it is sufficient to work out the derivation for $l \geq 0$. To complete the derivation, we first evaluate $\sigma_{ev}[l]$ for $l \geq 0$.
- A careful examination of (24) reveals that $v[k-l]$ contains effects of only past $e[k]$. By definition of WN, therefore, $\sigma_{ev}[l] = 0, l > 0$.

ACF of an AR(1) process ...contd.

To obtain $\sigma_{ev}[0]$, multiply both sides of (24) with $e[k]$ and take expectations on both sides to yield,

$$\begin{aligned} E(e[k]v[k]) &= -d_1 E(e[k]v[k-1]) + E(e[k]e[k]) \\ &= \sigma_e^2 \end{aligned}$$

using the same arguments as above. Thus, we have the following set of equations

$$\begin{aligned} \sigma_{vv}[0] &= -d_1\sigma_{vv}[-1] + \sigma_{ev}[0] \\ &= -d_1\sigma_{vv}[1] + \sigma_e^2 \\ \sigma_{vv}[1] &= -d_1\sigma_{vv}[0] \end{aligned}$$

ACF of an AR(1) process ...contd.

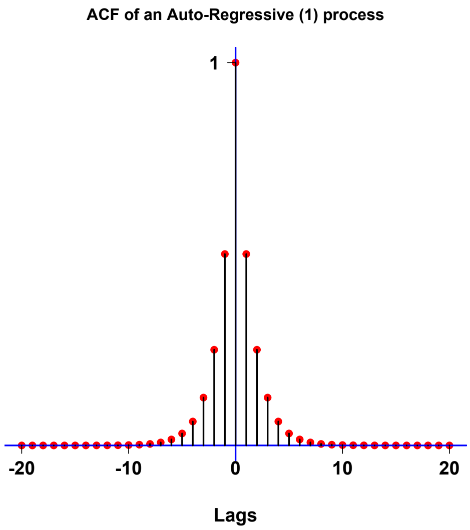
Solving equations for $\sigma_{vv}[0]$ and $\sigma_{vv}[1]$ simultaneously gives

$$\boxed{\begin{aligned} \sigma_{vv}[0] &= \frac{\sigma_e^2}{1 - d_1^2} \\ \rho_{vv}[l] &= (-d_1)^{|l|} \;\; \forall \;\; |l| \geq 1 \end{aligned}} \tag{25}$$

ACF of an AR(1) process ...contd.

- Shown adjacent is the plot of the ACF of an AR(1) process with $d_1 = -0.5$
- In general wheenver $|d_1| < 1$, we have that

The ACF of an AR(1) process exhibits exponential decay



Summary

- The ACF measures linear dependencies between observations of a time-series
- For a stationary process, the ACF is a symmetric function
- The ACF coefficients at any lag determine the optimal linear model for $x[k]$ in terms of its past.
- For an MA(M) process, the ACF abruptly vanishes after lags $|l| > M$.
- For an AR(P) process, the ACF dies down only exponentially.
  - The ACF satisfies the same difference equation as the random process itself